Perturbation biology models predict c-Myc as an effective co-target in RAF inhibitor resistant melanoma cells

Anil Korkut[1*], Weiqing Wang[1*], Emek Demir[1], Bülent Arman Aksoy[1,6], Xiaohong Jing[1], Evan Molinelli[1], Özgün Babur[1], Debra Bemis[1], David B. Solit[2,3], Christine Pratilas[4,5], Chris Sander[1*]

1. Computational Biology Center; 2. Human Oncology and Pathogenesis Program; Departments of 3. Medicine, 4. Pediatrics; 5. Program in Molecular Pharmacology and Chemistry, Memorial Sloan Kettering Cancer Center; 6. Tri-Institutional Training Program in Computational Biology & Medicine, New York, NY 10065, USA.

*Correspondence: bpmel@cbio.mskcc.org (will reach all principal authors)

Supplementary Information

1. Computational Methods

1.1. Automated extraction and reduction of prior information from signaling databases

Requirements for prior information. There is a growing number of signaling databases (Bader et al., 2006) that capture pathway information in high detail. Compared to the interaction networks with relatively low level of detail, pathway databases contain information about the phosphorylation states of proteins and interactions specific to these states. Although this corpus is highly valuable for phospho-proteomic analysis, it remains mostly untapped. To utilize this knowledge, one needs algorithms to map the experimentally measured proteins and phospho-proteins to their pathway database equivalents, find connections between them and reduce the outcome to a format that can be used for inference and analysis.

The Algorithm We developed an algorithm and a software tool (Prior Extraction and Reduction Algorithm, PERA) to automatically extract prior information from multiple signaling databases in the BioPAX (Demir et al., 2010) format and generate a prior information network. PERA takes a list of (phospho)proteins identified by their HGNC symbols (*e.g.* AKT1), phosphorylation sites (*e.g.* pS473) and their molecular status (*i.e.*, activating or inhibitory phosphorylation, total concentration) as input and then finds directed signaling paths between these entities. These paths are then reduced to directed interactions between signaling molecules represented in a Simple Interaction Format (SIF).

The PERA framework. The prior information is extracted from components of the Pathway Commons 2 database in three steps: First, using the paths-between

graph query algorithm (Dogrusoz et al., 2009), PERA generates a sub-graph (i.e., graph-of-interest) of Pathway Commons, which contains all the input proteins and all known connections within their first neighborhoods. Second, using the phosphorylation and activity state information, input entities are mapped to the corresponding protein states in the graph-of-interest. During this mapping step, protein states that do not match with either the corresponding annotation for phosphorylation or activity state are filtered out. Phosphorylation site mismatches up to 6 residues are tolerated during the filtering step to account for phosphorylation site ambiguities due to either database curation errors or cross-organism annotations. Third, paths that result in the addition or removal of a phopho-group at a phosphorylation site are extracted and mapped to phosphoprotein nodes. For total protein nodes, all non-phosphoprotein specific, directed signaling paths are included. For this application, the maximum allowable graph-query distance limit was set to 1 and only the Reactome (Matthews et al., 2009) and NCI-Nature PID pathway (Schaefer et al., 2009) data resources were used. Although we limited ourselves to short path distances and two pathway databases that we were most familiar with, PERA can be applied to extract information from any pathway database that exports to BioPAX and can be configured for searching paths of arbitrary length. The PERA software is available at http://bit.ly/bp_prior as free software under LGPL 3.0.

Key considerations for extracting pathway information: Two major semantic issues need to be taken into account when extracting state-specific pathway information from databases. First semantic issue stems from the ambiguities in mapping proteins with multiple phosphorylation sites. When a protein with multiple phosphorylation sites is experimentally profiled with an antibody, which recognizes only a single phosphorylation site, the antibody will actually bind to a heterogeneous mixture of phospho-states provided that the epitope is phosphorylated (e.g., anti-AKTpS473 Ab may bind to both AKTpS473 and AKTpS473_pT308 but not to AKTpT308). For proteins with multiple observed phosphorylation sites, this might lead to semantic conflicts since a double phosphorylated node should be mapped to both observations (i.e. single and double phosphorylated states). We included an optional “strict” mapping scheme to map only the phosphoproteins that exactly match the observed node - in our case always single phosphoproteins (e.g., the epitope of anti-AKTpS473 Ab is mapped to AKTpS473 but not to AKTpS473_pT308). Since our inference algorithm is much more tolerant of missing interactions compared to false interactions, we opted to use this flag for this application.

The second semantic issue stems from the fact that pathway databases are often curated from multiple independent studies, spanning multiple cellular states, cell and tissue types and even multiple model organisms. As a result, they are a superimposition of possible interactions over a wide range of spatiotemporal and genetic contexts. On the other hand databases cover only a subset of all possible contexts. In our case, we expect only a subset of the interactions in the pathway databases to be active in our cell lines and cover only a subset of the interactions

that we observe. This observation necessitates incorporating prior information as "soft" restraints for network inference. For this purpose, we devised a modified cost function, which includes a term for prior information (See below. Also see (Molinelli, Korkut, Wang et al., 2013)).

1.2. Modified cost function for network inference with prior information

In order to generate predictive network models of signaling, we have quantified the cost of W by an objective cost function C(W) that penalizes (a) discrepancies between predicted (x_i^{μ}) and experimental ($x_i^{\mu^*}$) values of the system variables at a time points $\{t_l\}$ in condition μ, and (b) the number of non-zero interactions with an L0 norm (Molinelli, Korkut, Wang et al., 2013; Nelander et al., 2008). Here we have also incorporated prior information to construct network models with higher predictive power even if the models are constructed with limited experimental data. The newly introduced third term in the cost function accounts for the prize introduced when the inferred w_{ij} conforms to the prior information (w_{ij}^{prior}). The modified cost function with prior information term is formulated as in equation S1.

Equation S1. Modified Cost function

$$C(W) = C^{SSE}(W) + C^{complexity}(W) + C^{prior}(W) \qquad 1.a$$

$$C(W) = \beta \sum_{l}^{L} \sum_{i}^{N} \sum_{\mu}^{M} \left(x_i^{\mu}(t_l) - x_i^{\mu^*}(t_l) \right)^2 + \lambda \sum_{i}^{N} \sum_{j \neq i}^{N} \delta(w_{ij}) + \sum_{i}^{N} \sum_{j \neq i}^{N} \eta(w_{ij}) \qquad 1.b$$

$$\delta(w_{ij}) = 1 \quad \text{if } w_{ij} \neq 0 \qquad 1.c$$

$$\delta(w_{ij}) = 0 \quad \text{if } w_{ij} = 0$$

In Equation S1, the first term accounts for the discrepancies between predicted (x_i^{μ}) and experimental ($x_i^{\mu^*}$) values of the system variables at a time point $\{t_l\}$ in condition μ, for a particular network configuration (**W=** $\{w_{ij}\}$). The second term is the complexity factor with an L0 norm, which reduces the number of nonzero interactions in a network configuration and ensures that resulting network models are sparse (Molinelli et al., 2013). The final term, η(**W**) is the cumulative prior cost function for **W=**$\{w_{ij}\}$. $\eta(w_{ij}= \omega)$ has a negative real value and reduces the overall model cost if the inferred ω conforms to the prior information. Here we will formulate the newly introduced prior information term in the modified cost function.

Generalized prior information prize function. In order to derive the prior information prize (PIP) function, we first assume the value of inferred w_{ij} is normally distributed around an expected prior value, $E[w_{ij}^{prior}]$ with a standard deviation, σ_{ij}^{prior}. The prior probability model for each interaction in the prior

network is formulized using the normal distribution probability density function around $E[w_{ij}^{prior}]$ as in equation S2.

Equation S2. Probability model of prior observations

$$P^{prior}(w_{ij}) = \frac{1}{\sigma_{ij}^{prior}\sqrt{2\pi}} \exp\left(-\frac{(w_{ij} - E[w_{ij}^{prior}])^2}{2(\sigma_{ij}^{prior})^2}\right)$$

The equation S2 implies that the probability of inferring a particular w_{ij} value and consequently, the prize introduced to the cost function increase as w_{ij} approaches $E[w_{ij}^{prior}]$. We calculate the prize for the fit between the inferred w_{ij} and the $E[W_{ij}^{prior}]$ with an inverse Boltzmann conversion of the probability model (Jaynes, 1957).

Equation S3. Error model of individual prior observations

$$\eta(w_{ij}) = -\kappa_{ij} \ln\left(P^{prior}(w_{ij})\right) \qquad 3.a$$

$$\eta(w_{ij}) = -\kappa_{ij} \ln\left(\frac{1}{\sigma_{ij}^{prior}\sqrt{2\pi}} \exp\left(-\frac{(w_{ij} - E[w_{ij}^{prior}])^2}{2(\sigma_{ij}^{prior})^2}\right)\right) \qquad 3.b$$

$$\eta(w_{ij}) = -\kappa_{ij} \ln\left(\frac{1}{\sigma_{ij}^{prior}\sqrt{2\pi}}\right) + \kappa_{ij}\left(\frac{(w_{ij} - E[w_{ij}^{prior}])^2}{2(\sigma_{ij}^{prior})^2}\right) \qquad 3.c$$

The first term introduces the maximum possible prize for each interaction that is represented in the prior information network. Second term penalizes the discrepancies between the w_{ij} and the $E[w_{ij}^{prior}]$. The penalty for the discrepancy is zero when $w_{ij} = E[w_{ij}^{prior}]$ and the prior prize, $\eta(w_{ij})$ assumes the highest possible value. $\kappa_{ij}(\ln(1/(\sigma_{ij}^{prior}\sqrt{2\pi}))$ is a constant analogous to the inverse temperature in statistical physics. The constant κ_{ij} and the resulting $\eta(w_{ij})$ is zero if the interaction between nodes i and j is not contained in the prior network model. By summing over all i and j, the cumulative prior prize for all $W=\{w_{ij}\}$ becomes:

Equation S4. Cumulative error model of prior observations

$$\eta(W) = \sum_{i}^{N}\sum_{j}^{N} \eta(w_{ij}) \qquad 4.a$$

$$\eta(W) = \sum_{i}^{N}\sum_{j}^{N} \left(-\kappa_{ij} \ln\left(\frac{1}{\sigma_{ij}^{prior}\sqrt{2\pi}}\right) + \kappa_{ij}\left(\frac{(w_{ij} - E[w_{ij}^{prior}])^2}{2(\sigma_{ij}^{prior})^2}\right)\right) \qquad 4.b$$

Even though the current PIP function is derived with the normal distribution assumption, a variety of alternative distributions such as bimodal distribution can

be fitted in Equation S2. Such efforts will lead to a variety of custom designed PIP functions when necessary.

Estimation of $E[w_{ij}^{prior}]$ and σ_{ij}^{prior}

An important requirement for the implementation of equation S4 is the accurate estimation of $E[w_{ij}^{prior}]$ and σ_{ij}^{prior}. Here, we propose two estimation strategies that are alternative to systematic extraction of priors from databases.

Priors from experiments. Our first strategy relies on use of biochemical measurements with multiple biological replicates to quantify the influence of node j on node I (w_{ij}). The $E(w_{ij}^{prior})$ and σ_{ij}^{prior} can be estimated from the mean value ($<w_{ij}^{prior}>$) and standard deviation (σ_{ij}) of the readouts from biological replicates in such measurements. Depending on the nature of the interaction under study, a variety of different biochemical methods can be utilized. These methods include but are not limited to enzyme activity assays, quantification of protein-protein interactions (Selbach and Mann, 2006), or additional RPPA readouts under specific perturbation conditions.

Priors from network models. As we build network models of signaling in diverse biological contexts using perturbation data and inference algorithms, a vast amount of information on signaling properties will become available. Consequently, we will be able to estimate the $E(w_{ij}^{prior})$ and σ_{ij}^{prior} for newly inferred models from available network models. Previously inferred $\{w_{ij}\}$ values and standard deviation in BP-generated probability distributions will serve as a basis to estimate the $E(w_{ij}^{prior})$ and σ_{ij}^{prior}, respectively.

The simplified PIP function

The current form of the prior information is limited to a set of binary interactions due to the qualitative nature of the databases, from which we extract the information. For instance, the prior information stored in the databases may correspond to activating ($w_{ij}>0$), inhibitory ($w_{ij}<0$), or generic ($w_{ij}\neq0$) interactions. In such situations, w_{ij}^{prior} has a uniform distribution within the defined boundary. Thus, the $E[w_{ij}^{prior}]$ can be set as equal to any value within the interval, $[w_{ij}^{min\text{-}prior}, w_{ij}^{max\text{-}prior}]$. For example a prior for a positive interaction between nodes i and j is represented by an interval of $(0, w_{ij}^{max}]$, where w_{ij}^{max} is the maximum allowed edge strength in the inference scheme. Consequently, the squared distance error in equation S3 vanishes and the prior prize assumes a fixed value over the defined prior interval.

Equation S5. Simplified error model of prior information

$$E[w_{ij}^{prior}] = w_{ij},\ \forall w_{ij} = \{w \in R \left| w_{ij}^{min-prior} \leq w \leq w_{ij}^{max-prior} \right.\}, \forall\ i,j \in N \quad 5.a$$

$$\kappa_{ij}\left(\frac{(w_{ij} - E[w_{ij}^{prior}])^2}{2(\sigma_{ij}^{prior})^2} = 0\right) \quad 5.b$$

$$\eta(w) = \sum_{i}^{N}\sum_{j}^{N}\left(-\kappa_{ij}\ln\left(\frac{1}{\sigma_{ij}^{prior}\sqrt{2\pi}}\right)\right) \quad 5.c$$

The binary nature of the interactions also implies a generic κ value ($\kappa=\kappa_{ij}(\ln(1/(\sigma_{ij}^{prior}\sqrt{2\pi}))$, $\kappa_{ij}>0$, $\forall$ i,j $\in$ N) for all interactions represented in prior information network and κ=0 for all other situations. The resulting prior information term in the cost function is a step function as shown in the methods function.

Equation S6. Simplified Error model of individual prior observations

$\eta(w_{ij} \neq 0) = -\kappa$ *and* $\eta(w_{ij} = 0) = 0$ *Generic prior information* ($w_{ij} \neq 0$)

$\eta(w_{ij} > 0) = -\kappa$ *and* $\eta(w_{ij} \leq 0) = 0$ *Prior information for an activating interaction* ($w_{ij} > 0$)

$\eta(w_{ij} < 0) = -\kappa$ *and* $\eta(w_{ij} \geq 0) = 0$ *Prior information for an inhibitory interaction* ($w_{ij} < 0$)

$\eta(w_{ij}) = 0$ *No Prior exists for* w_{ij}

For each interaction in the prior information network, $\eta(w_{ij})$ has a negative value. Thus, the prior information introduces a prize in the overall cost function. The formulation and implementation into the BP equations (see main text) create a soft-restraint for prior information extracted from signaling databases. In our probabilistic inference framework, a prior information is accepted as an edge ($w_{ij} \neq 0$) only if the prize introduced can override any discrepancy between predicted (x_i^{μ}) and experimental ($x_i^{\mu^*}$) values of the system variables. Thus the context specific character of inferred models is preserved when the cost function in equation S1 is used.

1.3. Network inference

The computational/experimental network modeling procedure is described in Figure 1. The theoretical details of the BP-based network inference is described elsewhere (Molinelli, Korkut, Wang et al. 2013). The BP-guided decimation algorithm is described in Figure S1. The mean-field approximation in belief propagation inference algorithm is summarized below.

Gaussian integration of cavity parameter update. Information from a sufficiently large set of non-cavity constraints and parameters are integrated to update the cavity parameters. As a corollary to the central limit theorem, we can assume that the mean-field effect of non-cavity parameters on the cavity update follows a Gaussian distribution. The mean and variance of the Gaussian

distribution are the means and variances of the global distributions given by individual distributions of non-cavity $P^{\mu}(w_{ij})$ $\forall j \neq k$. Thus, we replace the sum over multivariate configurations (Equation 2.B in main text) of all non-cavity parameters with a single Gaussian integration over the mean-field of non-cavity configurations.

Equation S7. Gaussian approximation to the cavity parameter update

$$\rho^{\mu}(w_{ik}=\omega)=\frac{1}{Z_{ik}}\frac{1}{\sqrt{2\pi\Delta_k^{\mu}}}\int_{-\infty}^{\infty} e^{-\beta\left(x_i^{\mu^*}-\phi(s_k^{\mu}+w_{ik}x_k^{\mu^*})\right)^2} e^{-\frac{(\overline{s_k^{\mu}}-s_k^{\mu})^2}{2\Delta_k^{\mu}}} ds_k^{\mu} \quad \forall\omega\in\Omega \qquad 7.a$$

$$\overline{s_k^{\mu}}=\sum_{j\neq i,k}^{N}\overline{w_{ij}}x_j^{\mu^*}+u_i^{\mu} \qquad 7.b$$

$$\Delta_k^{\mu}=\sum_{j\neq i,k}^{N}\left(\overline{w_{ij}}^2-\overline{w_{ij}^2}\right)(x_j^{\mu^*})^2 \qquad 7.c$$

$$\overline{w_{ij}}=\sum_{\omega}\omega P^{\mu}(w_{ij}=\omega) \qquad 7.d$$

Note that, BP parameters (temperature factor, β=2, complexity factor, λ=5) are adapted for all computations after a systematical analysis of both parameters described in Molinelli, Korkut, Wang et al. Prior information parameter, κ is set as (κ= λ =5) to ensure prior information and sparsity restraints in the models balance each other. When prior information is used to construct the models, the complexity term, λ is set to 2.5 for phenotypic nodes in order to approximately match the complexity around the phenotypic nodes and rest of the networks. When the parameter space is sufficiently large, this approximation significantly reduces the computational complexity of the problem without loss of overall accuracy. By means of the iterative approach with the mean-field Gaussian approximation, the time-complexity of the problem is strongly reduced and the obstacles imposed by combinatorial complexity are circumvented.

2. Experimental methods

Cell cultures and Perturbation experiments. In perturbation experiments, the drugs are applied to cell cultures after SkMel133 cells are grown to about 40% confluence in complete RPMI-1640 medium (10% heat-inactivated fetal bovine serum, 100 units/ml each of penicillin and streptomycin, and incubated at 37°C in 5% CO_2) in 6-well plates. 24 hours after drug administration, the perturbed cells are harvested. In control experiments (i.e. no-drug condition), cells are treated with the DMSO drug vehicle for 24 hours.

Reverse Phase Protein Arrays. Proteomic response profiles are measured using reverse phase protein arrays (MD Anderson Cancer Center RPPA Core Facility) (Tibes et al., 2006). For sample preparation, cell pellets are lysed in RPPA buffer (1% Triton X-100, 50mM HEPES pH7.4, 150mM NaCl, 1.5mM

MgCl2 , 1mM EGTA, 100mM NaF, 10mM Na pyrophosphate, 1mM Na3VO4, 10% glycerol, freshly added protease and phosphatase inhibitors). After cells are lysed, total protein concentrations in cell lysates are measured with the Bradford assay and the final protein concentrations are adjusted to 1-1.5mg/ml. Samples are boiled in SDS sample buffer without bromophenol blue for 5 minutes at 95°C. Cell lysates are spotted on nitrocellulose-coated slides as described in (Tibes et al., 2006). Antibody staining intensities are quantified using the MicroVigene automated RPPA module (VigeneTech, Inc.) and the standard RPPA protein concentration normalization procedure (Neeley et al., 2009) is followed. The proteomic readouts are log normalized with respect to the corresponding untreated condition readouts. We have eliminated those readouts with intra or inter slide coefficient of variation (CV) > 0.15 (I.e., low reproducibility) and low degree of staining by antibodies. 100 proteomic entities are chosen for further analysis. Next, we excluded the proteomic entities that do not respond to at least a single perturbation condition from the network models. This has lead to highly reliable response measurements on a final set of 82 proteomic entities, which have been used in the network models.

Quantitative phenotypic assays. All phenotypic measurements are made in perturbation conditions identical to those in proteomic measurements. Cell viability is measured 72 hours after perturbation using the Resazurin assay. Each well is treated with the Resazurin (Sigma-Aldrich, Catalog # R7017) at a final concentration of 44 **μ**M for 1 hour. The fluorescent signals are measured at 530 nm excitation wavelength and 590 nm emission wavelength. Standard curves of cell numbers are used to calculate the cell numbers in each sample. The cell cycle progression phenotypes are measured with flow cytometry analysis. Cells are seeded in 6-well plates. 72 hours after drug application, cells are harvested by trypsinization, including both suspended and adherent fractions, and washed in cold PBS. Cell nuclei are prepared as described in (Nusse and Kramer, 1984) and cell cycle distribution was determined by flow cytometric analysis of DNA content using red fluorescence of 488 nm excited ethidium bromide-stained nuclei. The percentage of cells in the G1, G2/M, and S phases and sub-G1 fraction are recorded.

Supplementary References:

Bader, G.D., Cary, M.P., and Sander, C. (2006). Pathguide: a Pathway Resource List. Nucleic Acids Research *34*, D504-D506.

Demir, E., Cary, M.P., Paley, S., Fukuda, K., Lemer, C., Vastrik, I., Wu, G.N., D'Eustachio, P., Schaefer, C., Luciano, J.*, et al.* (2010). The BioPAX community standard for pathway data sharing. Nature Biotechnology *28*, 935-942.

Dogrusoz, U., Cetintas, A., Demir, E., and Babur, O. (2009). Algorithms for effective querying of compound graph-based pathway databases. Bmc Bioinformatics *10*.

Jaynes, E.T. (1957). Information Theory and Statistical Mechanics. Physical Review *106*, 620-630.

Matthews, L., Gopinath, G., Gillespie, M., Caudy, M., Croft, D., de Bono, B., Garapati, P., Hemish, J., Hermjakob, H., Jassal, B.*, et al.* (2009). Reactome knowledgebase of human biological pathways and processes. Nucleic acids research *37*, D619-622.
Molinelli, E.J., Korkut, A., Wang, W., Miller, M.L., Gauthier, N.P., Jing, X., Kaushik, P., He, Q., Mills, G., Solit, D.B.*, et al.* (2013). Perturbation biology: inferring signaling networks in cellular systems. PLoS computational biology *9*, e1003290.
Neeley, E.S., Kornblau, S.M., Coombes, K.R., and Baggerly, K.A. (2009). Variable slope normalization of reverse phase protein arrays. Bioinformatics *25*, 1384-1389.
Nelander, S., Wang, W., Nilsson, B., She, Q.B., Pratilas, C., Rosen, N., Gennemark, P., and Sander, C. (2008). Models from experiments: combinatorial drug perturbations of cancer cells. Molecular systems biology *4*, 216.
Nusse, M., and Kramer, J. (1984). Flow Cytometric Analysis of Micronuclei Found in Cells after Irradiation. Cytometry *5*, 20-25.
Schaefer, C.F., Anthony, K., Krupa, S., Buchoff, J., Day, M., Hannay, T., and Buetow, K.H. (2009). PID: the Pathway Interaction Database. Nucleic acids research *37*, D674-679.
Selbach, M., and Mann, M. (2006). Protein interaction screening by quantitative immunoprecipitation combined with knockdown (QUICK). Nature methods *3*, 981-983.
Tibes, R., Qiu, Y., Lu, Y., Hennessy, B., Andreeff, M., Mills, G.B., and Kornblau, S.M. (2006). Reverse phase protein array: validation of a novel proteomic technology and utility for analysis of primary leukemia specimens and hematopoietic stem cells. Molecular cancer therapeutics *5*, 2512-2521.

Supplementary Figures:

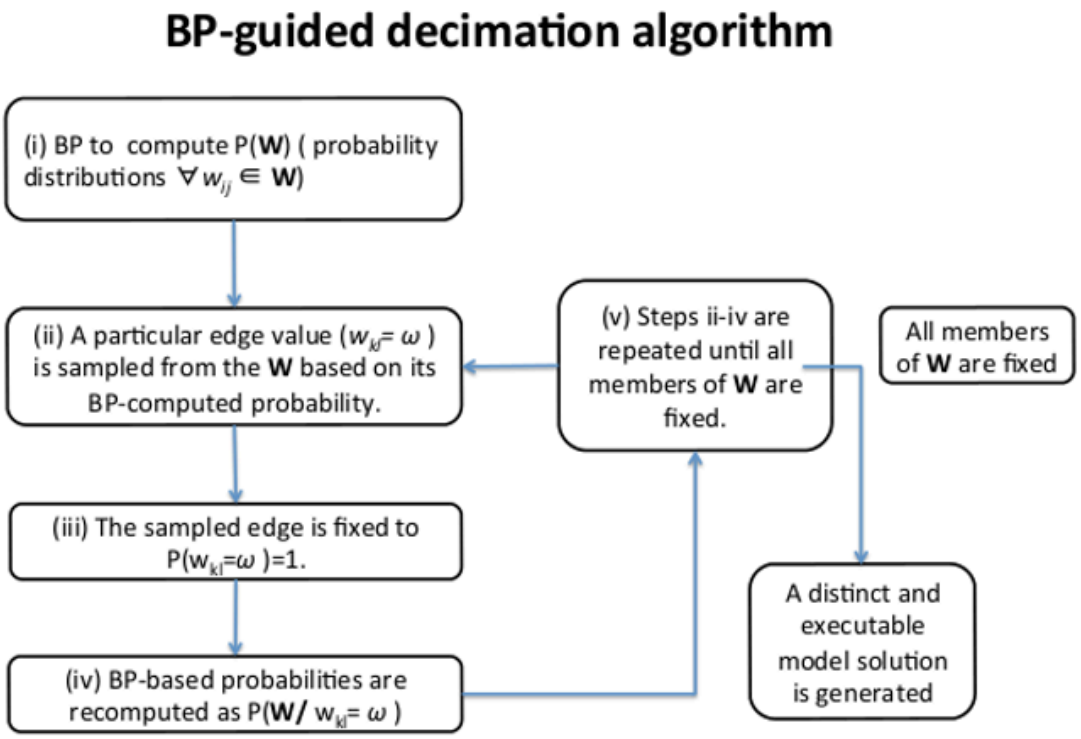
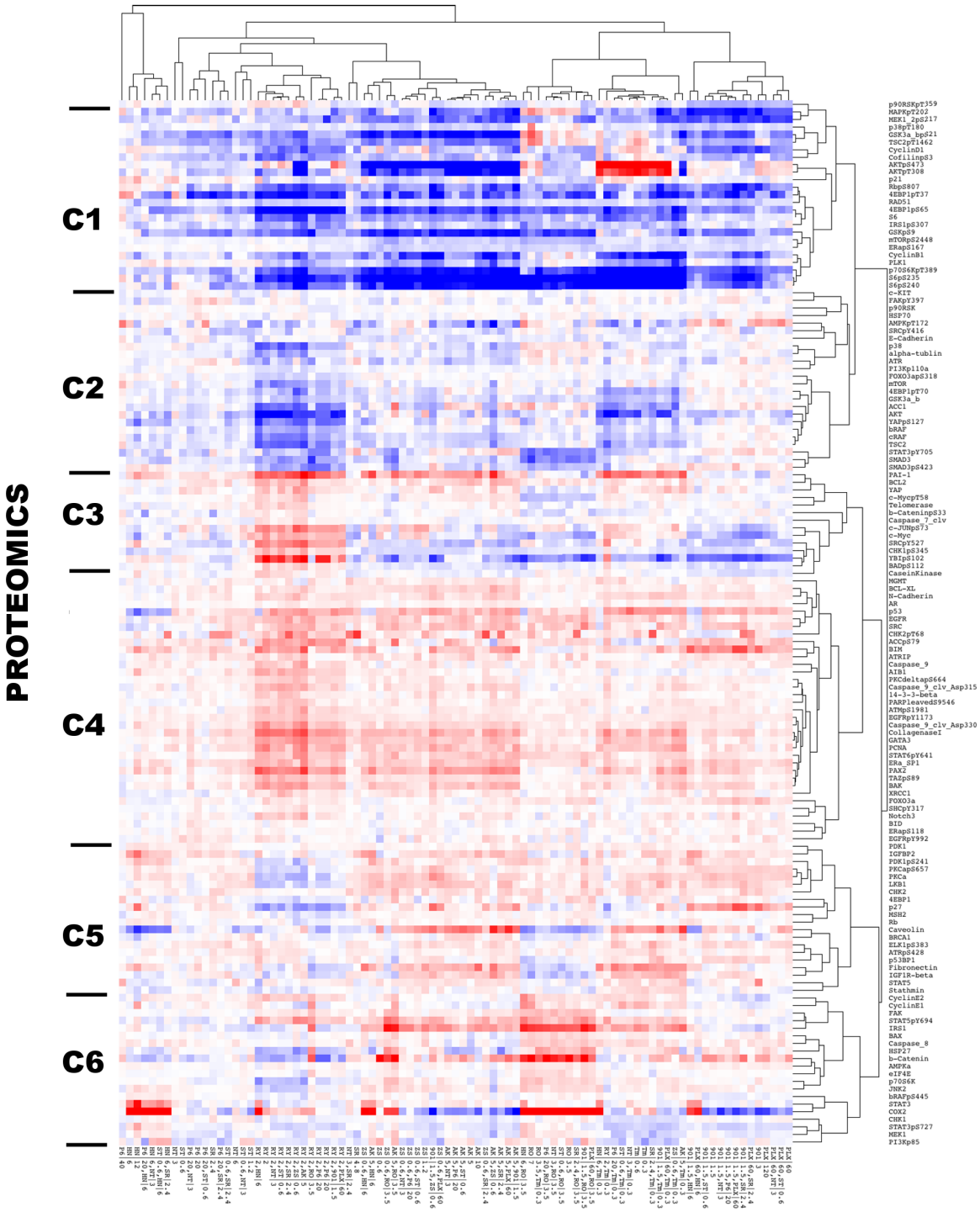


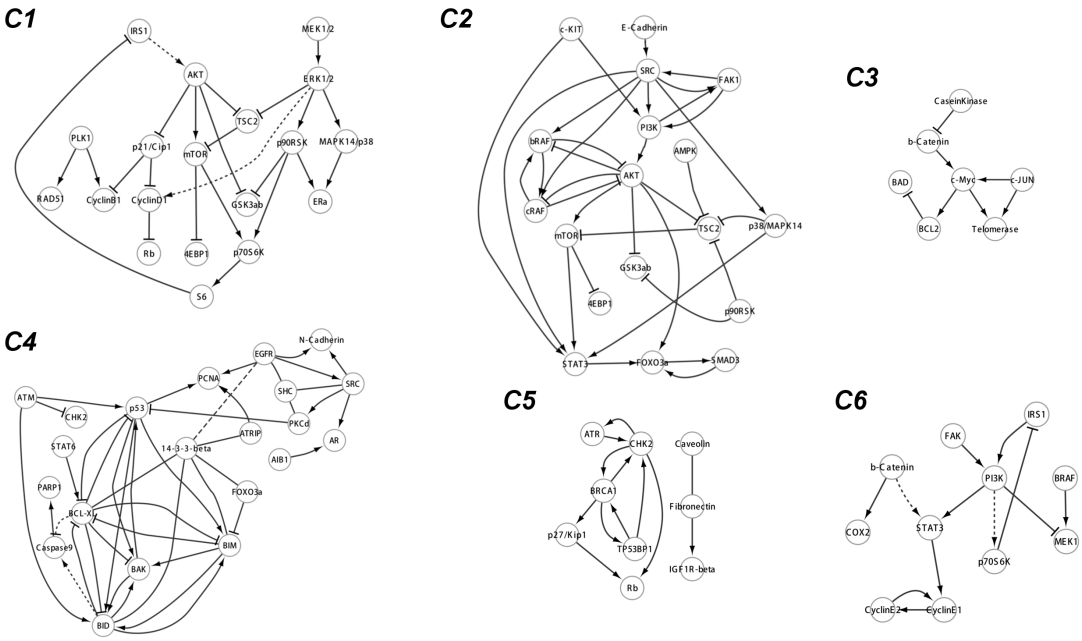
Figure S1 BP-guided decimation algorithm is used to construct executable, individual network model solutions from BP generated probability distributions for each edge strength value (wij). The algorithm chart depicts one round of BP-guided decimation to generate a single model solution. Consecutive runs of BP-guided decimation algorithm leads to construction of a network model solution ensemble.

A **CLUSTERING OF PROTEOMIC RESPONSE DATA**

PERTURBATIONS



B PATHWAY ANALYSIS DRIVEN BY PROTEOMIC CLUSTERS



c

Proteomic response to targeted drugs in SkMel133 cells

MEKi (1.5e-3uM)		MEKi (3e-3 uM)		AKTi (5uM)		AKTi (10uM)	
b-Catenin	1.346	BIM	1.510	Caveolin	1.428	Fibronectin	1.509
BIM	1.296	p27	1.502	PAI-1	1.283	PAI-1	1.496
p27	1.244	STAT5	1.396	PAX2	1.278	Caveolin	1.382
ACCpS79	1.235	AMPKpT172	1.306	IRS1	1.266	IRS1	1.346
PAX2	1.229	b-Catenin	1.270	Fibronectin	1.250	PAX2	1.270
AMPKpT172	1.224	p53	1.232	BAK	1.238	p53	1.245
ERa_SP1	1.215	FOXO3a	1.229	BCL2	1.234	IGFBP2	1.243
PKCa	1.212	AKTpS473	1.219	Collagenasel	1.226	CHK2pT68	1.217
p53	1.212	AKTpT308	1.219	STAT6pY641	1.212	STAT5pY694	1.192
Caveolin	1.180	CHK2	1.159	ATRIP	1.189	BCL2	1.182
GSKpS9	0.700	p70S6KpT389	0.632	TSC2pT1462	0.637	COX2	0.632
4EBP1pT37	0.687	MEK1/2pS220	0.610	4EBP1pS65	0.589	CyclinB1	0.623
YBIpS102	0.682	YBIpS102	0.593	AMPKpT172	0.578	MEK1/2pS221	0.589
CyclinD1	0.679	S6pS240	0.592	p70S6KpT389	0.539	GSKpS9	0.564
RbpS807	0.656	COX2	0.577	S6pS240	0.517	GSK3abpS21	0.550
GSK3abpS21	0.653	CyclinB1	0.569	GSKpS9	0.514	p70S6KpT389	0.505
p70S6KpT389	0.648	S6pS235	0.569	GSK3abpS21	0.440	S6pS240	0.418
S6pS235	0.614	CyclinD1	0.562	S6pS235	0.364	S6pS235	0.274
MEK1/2pS217	0.604	RbpS807	0.542	AKTpT308	0.250	AKTpT308	0.240
MAPKpT202	0.553	MAPKpT202	0.468	AKTpS473	0.173	AKTpS473	0.140

HDACi (6uM)		HDACi (12 uM)		MDM2i 3uM)		MDM2i (6uM)	
COX2	2.787	COX2	3.735	4EBP1pT37	1.271	STAT5	1.206
STAT3	1.412	STAT3	1.955	p21	1.197	ATR	1.133
IGFBP2	1.289	IGFBP2	1.487	S6pS240	1.171	p53	1.119
4EBP1pT37	1.272	p27	1.481	Collagenasel	1.156	CHK2pT68	1.114
AMPKpT172	1.162	BIM	1.344	p53	1.148	IGFBP2	1.104
Fibronectin	1.148	PAI-1	1.316	PCNA	1.130	PCNA	1.104
PAI-1	1.128	CyclinE1	1.287	AKT	1.115	COX2	1.104
BIM	1.114	STAT3pS727	1.274	mTOR	1.103	PAI-1	1.101
CyclinD1	1.108	Fibronectin	1.269	YBIpS102	1.083	Collagenasel	1.096
p21	1.087	p21	1.256	N-Cadherin	1.081	Caspase_8	1.090
CyclinB1	0.902	AKT	0.813	PI3Kp85	0.904	CofilinpS3	0.853
ACC1	0.891	IRS1pS307	0.809	MEK1/2pS239	0.898	S6pS235	0.845
4EBP1pS65	0.887	cRAF	0.790	E-Cadherin	0.901	GSKpS9	0.844
TSC2pT1462	0.887	S6pS235	0.775	AKTpS473	0.874	GSK3abpS21	0.842
MEK1/2pS234	0.861	RAD51	0.762	CofilinpS3	0.870	4EBP1pS65	0.828
SRC	0.851	c-Myc	0.734	p70S6KpT389	0.869	AKT	0.820
CHK1pS345	0.846	CyclinB1	0.729	TSC2pT1462	0.861	MEK1/2pS238	0.780
YAPpS127	0.831	YAPpS127	0.662	GSKpS9	0.825	S6pS240	0.801
p53	0.805	p53	0.644	GSK3abpS21	0.806	Caveolin	0.775
Caveolin	0.725	Caveolin	0.562	4EBP1pS65	0.746	4EBP1pT37	0.715

JAKi (0.02uM)		JAKi (0.04uM)		RAFi (0.06 uM)		RAFi (0.12 uM	
FAKpY397	1.161	AMPKpT172	1.445	b-Catenin	1.313	b-Catenin	1.294
IGFBP2	1.151	4EBP1pT37	1.343	p27	1.252	p27	1.282
Fibronectin	1.147	STAT5	1.240	Caveolin	1.219	BIM	1.238
YAP	1.120	mTOR	1.207	BIM	1.196	AMPKpT172	1.213
ACCpS79	1.116	b-Catenin	1.175	ERa_SP1	1.167	p53	1.207
HSP70	1.113	4EBP1	1.171	AMPKpT172	1.158	FOXO3a	1.171
MEK1/2pS237	1.106	MAPKpT202	1.167	PKCa	1.154	AKTpS473	1.170
EGFRpY992	1.111	S6pS235	1.158	ACCpS79	1.147	HSP27	1.156
BIM	1.104	p21	1.147	AKTpS473	1.134	AKTpT308	1.141
p27	1.100	S6	1.140	PCNA	1.131	PKCapS657	1.133
MAPKpT202	0.880	STAT5pY694	0.920	CofilinpS3	0.782	c-Myc	0.768
p70S6KpT389	0.876	BAK	0.917	RbpS807	0.780	p70S6KpT389	0.736
RbpS807	0.868	STAT3pS727	0.909	AKT	0.778	S6pS235	0.725
IRS1pS307	0.857	MSH2	0.903	GSKpS9	0.745	CyclinB1	0.720
CofilinpS3	0.843	Caveolin	0.893	YBIpS102	0.742	RbpS807	0.687
AKTpT308	0.839	CofilinpS3	0.884	COX2	0.731	YBIpS102	0.669
TSC2pT1462	0.780	p70S6KpT389	0.871	p70S6KpT389	0.712	MAPKpT202	0.636
GSKpS9	0.746	MEK1/2pS236	0.847	GSK3abpS21	0.692	CyclinD1	0.629
GSK3abpS21	0.695	AKTpT308	0.842	MAPKpT202	0.648	COX2	0.595
4EBP1pT37	0.603	STAT3pY705	0.816	MEK1/2pS218	0.582	MEK1/2pS217	0.352

C-Continued

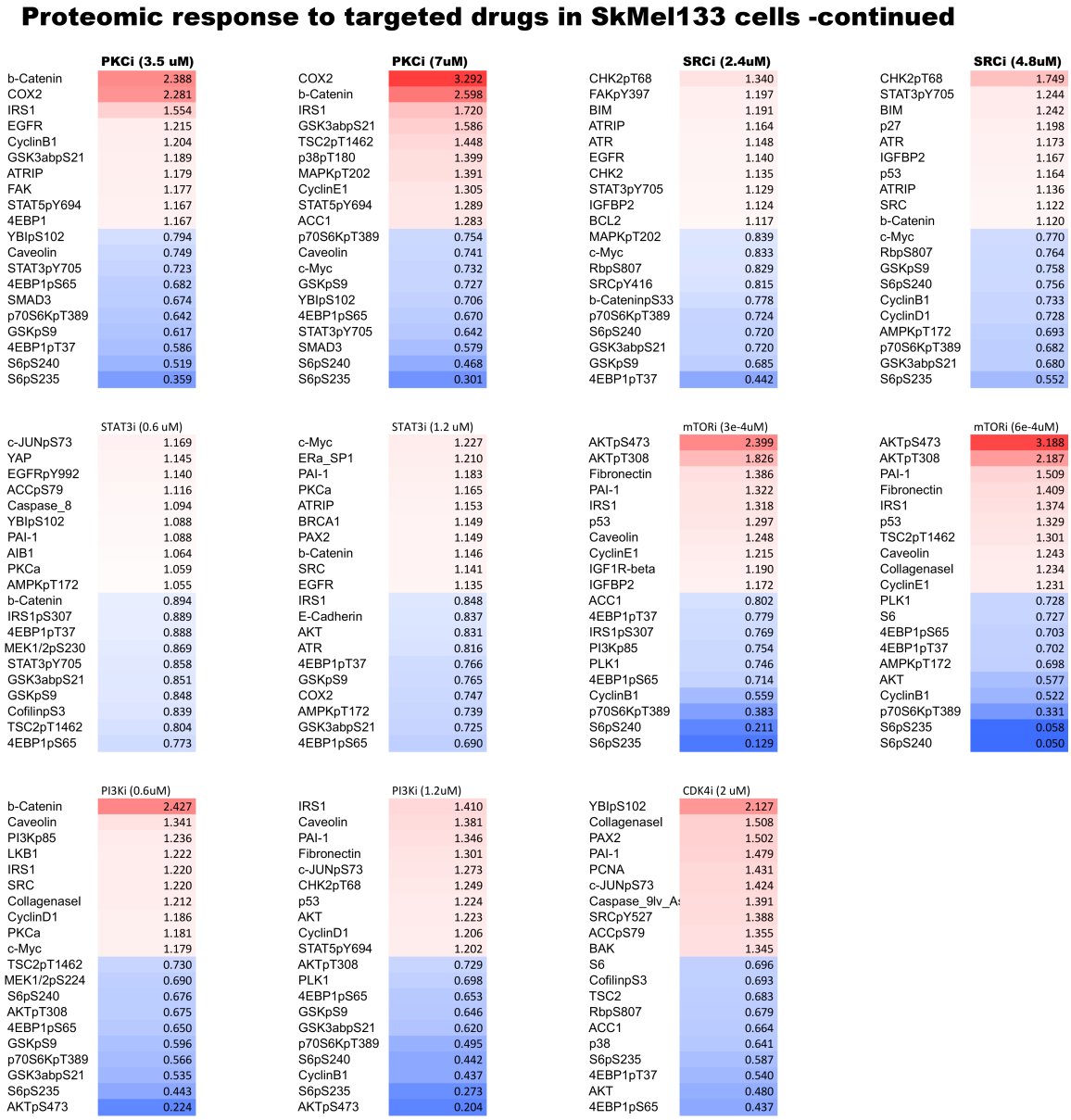
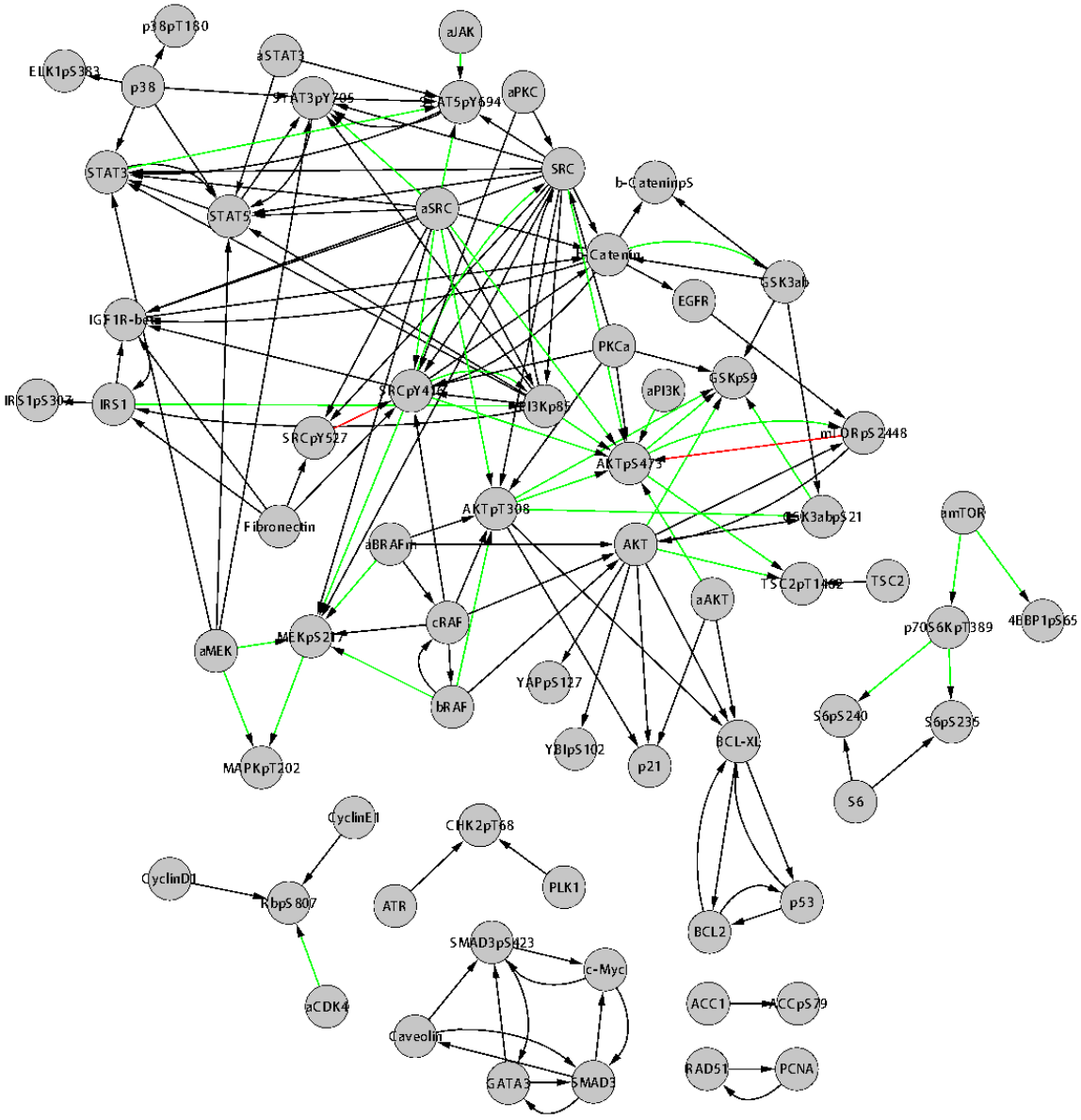


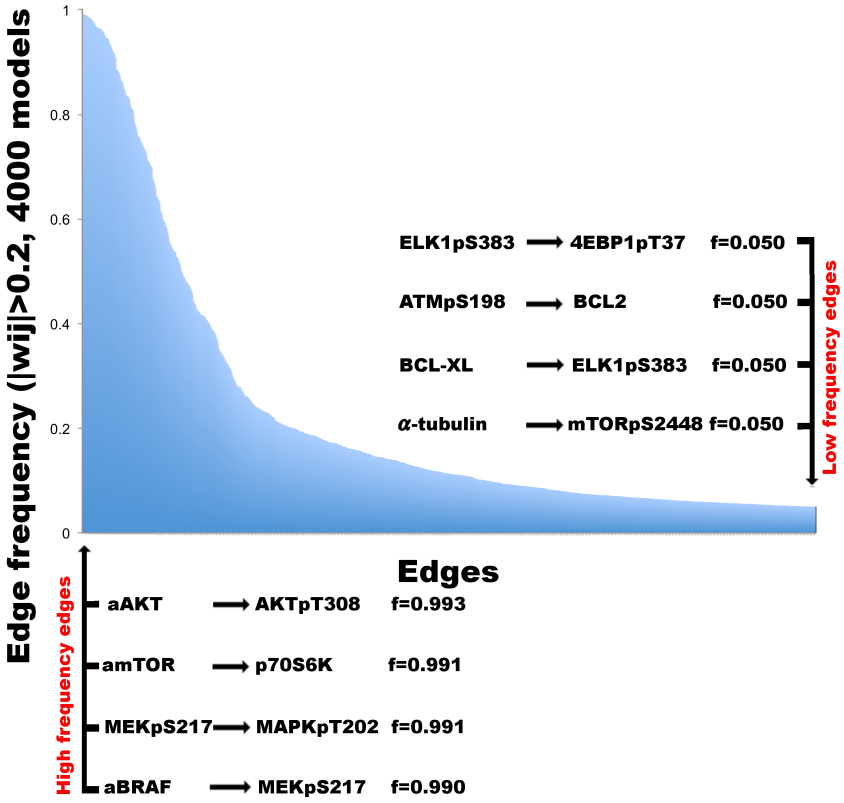
Figure S2 Clustering and pathway analysis of proteomics data. We characterized the proteomic signatures in the response data with a pathway analysis guided by hierarchical clustering. **A**. The two-way clustering analysis of the experimental response map reveals distinct proteomic signatures of response to drugs targeting different signaling pathways (See Figure 2 in main text). The Cluster software is used for the two way hierarchical clustering with correlation-

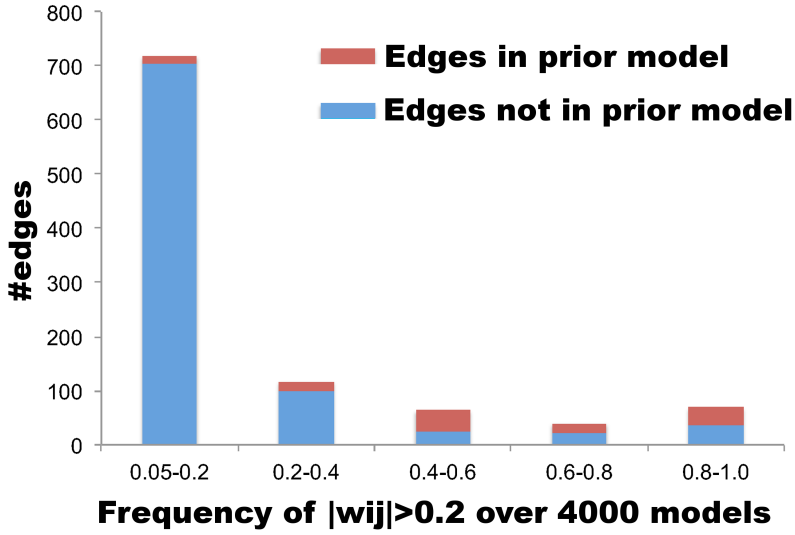
based distance metric and average-linkage method. We identified 3 major (C1, C2 and C4) and 3 minor (C3, C5, C6) proteomic response clusters. **B.** We extracted the signaling interactions between the proteomic entities that fell into each cluster. The signaling interactions within each proteomic cluster are extracted from Pathway Commons 2 database using the PERA algorithm. We simplified the pathway diagrams by removing the post-translational modifications and merging the nodes associated with identical genes (e.g., AKTpT308, AKTpS474 and AKT are merged to AKT node). The resulting diagrams displayed the known pathways associated with proteomic clusters, whose members gave similar response to targeted agents. The cluster guided pathway analysis suggests that functionally related proteins (i.e. proteins on the same or related pathways) are enriched in distinct clusters. The phospho-proteomic entities on MAPK and PI3K/AKT pathways are enriched in cluster 1 (C1) of the response map. The total protein measurements related to MAPK and PI3K/AKT pathways are enriched in a related but distinct cluster (C2). The level of various apoptotic proteins increase in response to most targeted drugs and form another distinct cluster (C4). Interestingly, EGFR (Both phospho and total level) and a set of EGFR related docking proteins (e.g., SHC, 14-3-3-β) are also enriched in C4 possibly due to the increase in the expression of those entities in response to targeted drugs such as MEKi and PI3Ki. The observed increase in EGFR level suggests a potential feedback loop that emerges from downstream elements of MAPK and PI3K/AKT pathways, which are targeted by multiple agents in this study. **C**. Proteomic response to single agent perturbations. Proteomic levels in SkMel133 cells change in response to targeted drugs as measured with RPPA. For each drug, we listed the top 10 decreasing and increasing phospho-proteins out of 138 proteomic entities profiled with RPPA (Figure S1A). The protein levels are normalized to no drug condition and given in linear space.

A Qualitative prior model



B **Statistical properties of the solution ensemble**





C The performance of pathway database driven priors over random priors

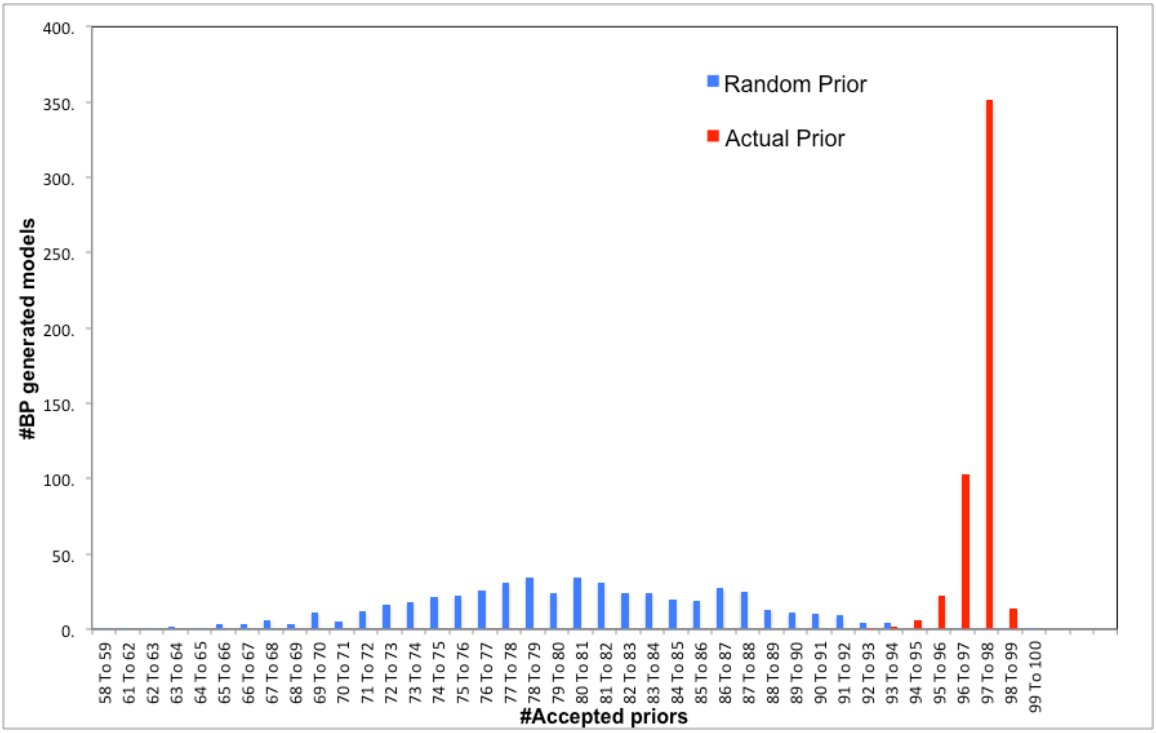
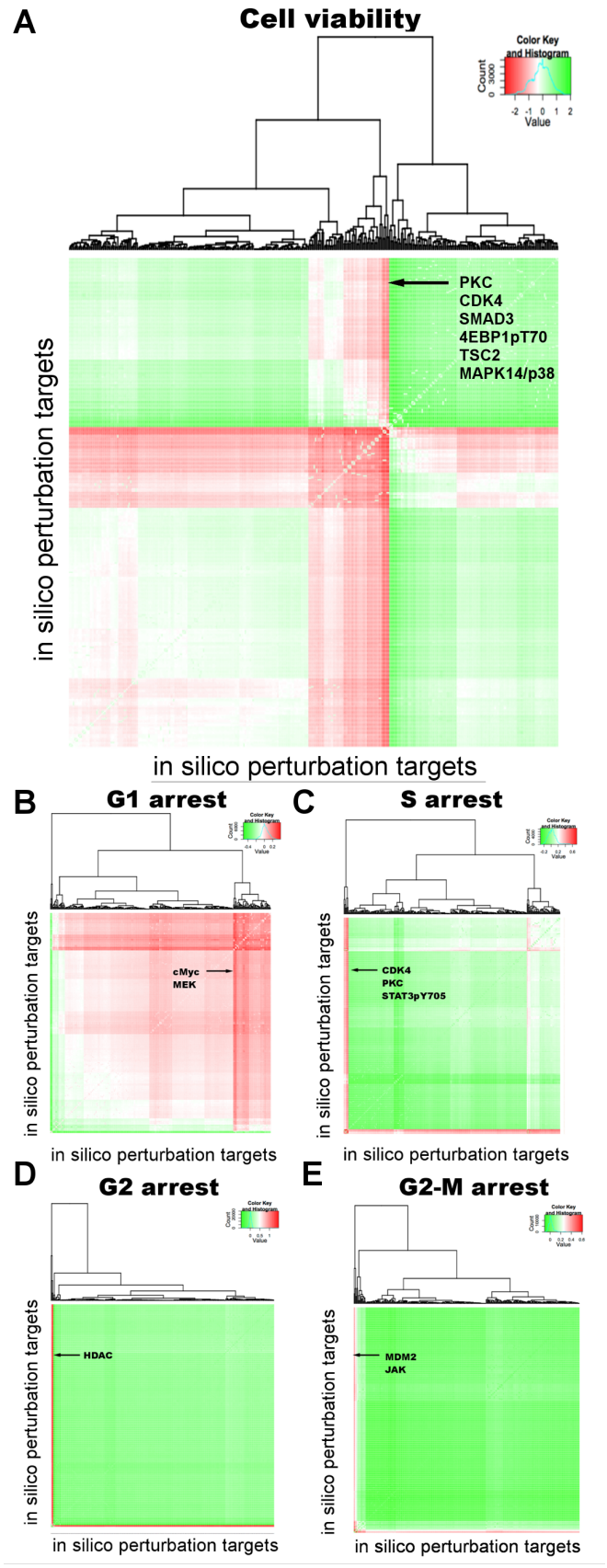


Figure S3. A. The prior model of signaling. The prior information model is extracted from Reactome and NCI PID using the PERA algorithm (See methods and supplementary methods for PERA algorithm and use of the prior model in inference). The prior model served as a bias in BP-based network inference. Green arrows represent priors for activating edges, red arrows represent inhibitory edges and black arrows represent generic edges (i.e. activating or inhibitory). **B. Distribution of edges in the solution ensemble.** In order to model cellular signaling and predict response, we computed 4000 signaling models and generated a solution ensemble (See main text). **A.** Edge frequencies ($f(w_{ij})$) in the solution ensemble. The y-axis represents the frequency values for nonzero edge strengths ($f(|w_{ij}|>0.2)$) in the ensemble. The X-axis represents the interactions ordered by their $f(|w_{ij}|>0.2)$. The long tail that corresponds to the edges with frequencies less than 0.05 is not displayed. A set of well reported interactions such as the coupling of AKT activity node to phospho-AKT and phospho-MEK to phosho-MAPK (ERK) are inferred with $|w_{ij}|> 0.2$ in >99% of the model solutions (left). Few examples of the rarely captured signaling interactions are given on right. B. The frequency distribution of nonzero edges ($|w_{ij}|>0.2$) has a bimodal character. A set of interactions with nonzero edge values is shared by nearly all of the inferred network models ($0.8 < f(|w_{ij}|> 0.2) < 1.0$). Such frequent edges form a stable network skeleton shared by majority of solutions. On the other hand, a set of possible interactions have non-zero edge strength (w_{ij}) values in around 50% of the solutions. Such interactions vary among

different model solutions and possibly account for the variations in system response predicted with different model solutions. The interactions in the prior model are particularly enriched in these two groups. A large fraction of potential interactions have very low ($0.05 < f(|w_{ij}|>0.2) < 0.20$) or near zero ($f(|w_{ij}|>0.2) < 0.05$) frequencies in the model ensemble. **C. Comparison of random vs. actual prior information.** We tested whether the database driven prior model conformed to the experimental data and prior information does not overly restrain the inferred models. For this purpose, we constructed and compared network models using the pathway driven and randomly generated prior restraints in BP-based modeling. We have tested 500 unique, randomly generated prior restraints each containing 154 interactions, equal to the number of interactions in the pathway driven prior model. For model construction, we first computed the probability distribution of edge values ($P(W_{ij})$) for each possible interaction. Next, we constructed the models by assigning the edge value (W_{ij}) for which BP-generated $P(W_{ij})$ is maximum. We compared the models generated with different prior models for their prior scores (i.e. number of edges, which were accepted by the algorithm and also contained in the prior model). The models generated using the database driven priors had significantly higher prior scores (ps) compared to randomly generated models ($p < 0.05$ Student's t-test for H0: $\mu_{ps}^{random} = \mu_{ps}^{database\text{-}driven}$). Note that, we performed 500 parallel BP runs with identical database driven prior models and experimental data. We observed a low degree of variation in the prior scores for the models with database driven priors. This is indicative of some solution degeneracy in the models inferred with database driven priors. In predictive network modeling, we accounted for the solution degeneracy by re-computing the $P(W_{ij})$ at the beginning of each BP-guided decimation step (See main text and methods) and ensuring that the final solution ensemble contains models that sample every possible degenerate solution.


A
Cell viability
Color Key and Histogram
Count
Value
PKC
CDK4
SMAD3
4EBP1pT70
TSC2
MAPK14/p38
in silico perturbation targets
in silico perturbation targets
B
G1 arrest
cMyc
MEK
in silico perturbation targets
in silico perturbation targets
C
S arrest
CDK4
PKC
STAT3pY705
in silico perturbation targets
in silico perturbation targets
D
G2 arrest
HDAC
in silico perturbation targets
in silico perturbation targets
E
G2-M arrest
MDM2
JAK
in silico perturbation targets
in silico perturbation targets

C

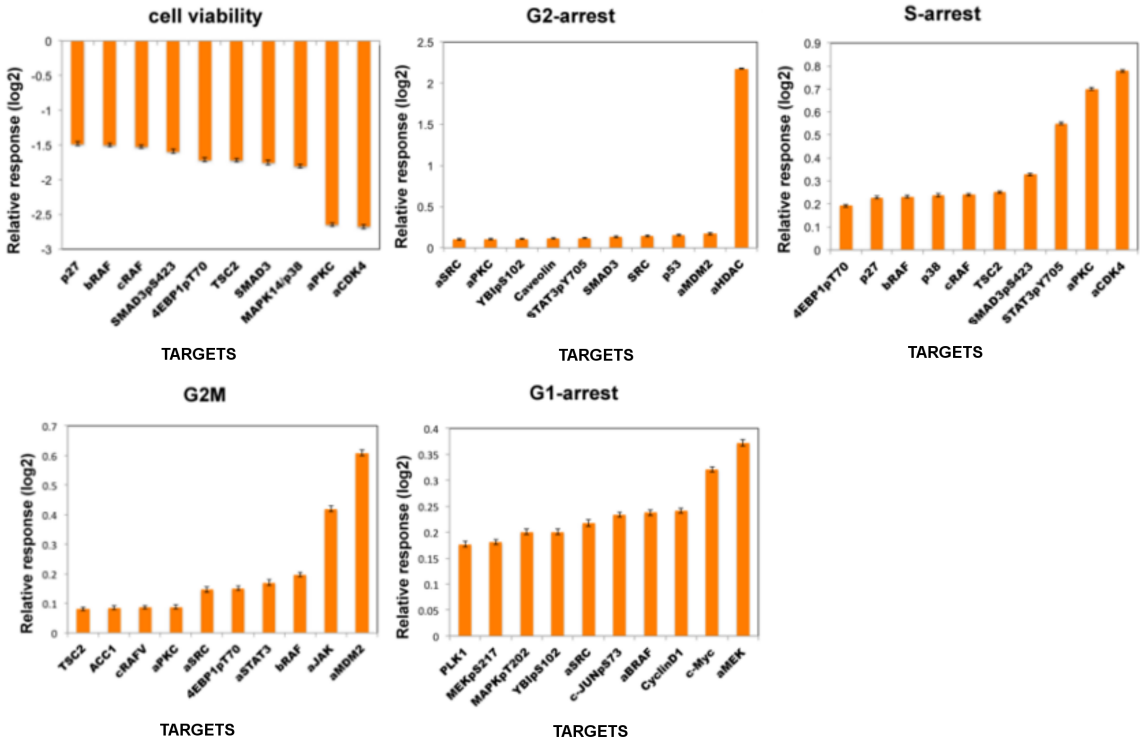
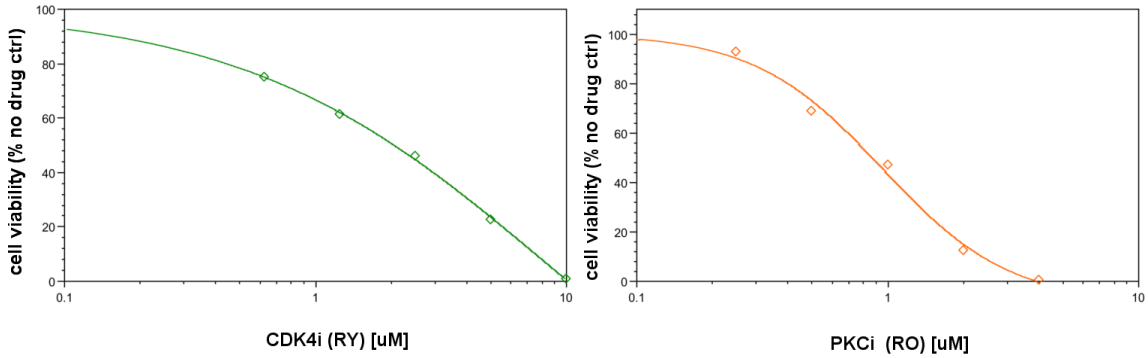
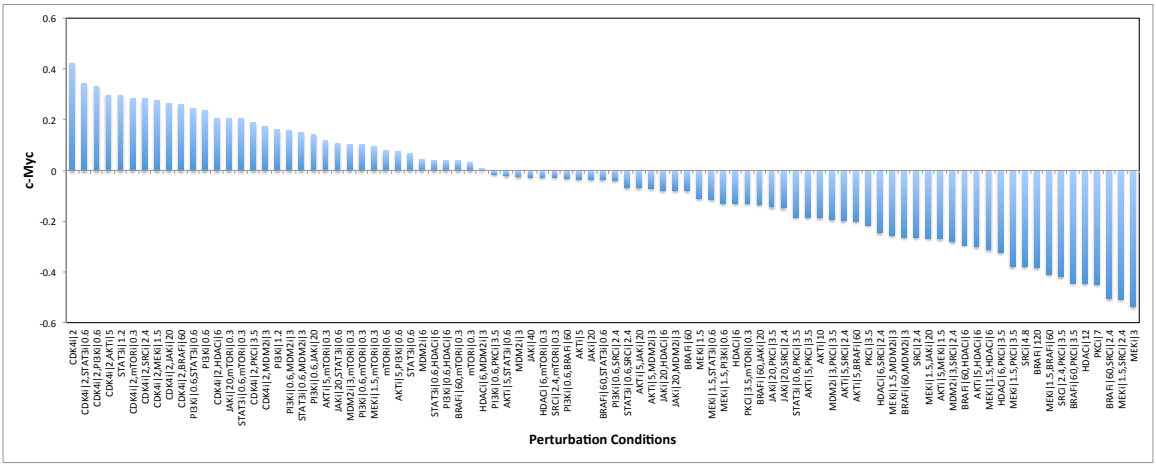


Figure S4 A. Prediction of phenotypic responses to in silico, combinatorial perturbations. We computationally predicted cell viability and cell cycle arrest (G1, S, G2 and G2M) in response to combinatorial perturbations of all proteomic nodes in network models (See Table 1 and Figure 5-6 in main text). In silico perturbations include paired combinations of 94 perturbations in 4 different concentrations (IC0-20-40-60-80 with respect to the basal level of a particular node) leading to more than 70000 unique perturbation conditions. Each in silico perturbation is applied to 4000 inferred network models. This leads to a simulation of ~28 million unique conditions for 99 different readouts, a throughput level inaccessible by experimental screening methods. The heatmaps display the phenotypic response landscape to combinatorial perturbations. Each data point on heatmaps is the response to a perturbation averaged over predictions from 4000 network models **A.** Cell viability response to perturbations. The desired response from a perturbation is reduction in cell viability. (Red: decreased cell viability, green: increased cell viability). **(B-E).** Same as in **A** but for cell cycle progression phenotypes. For cell cycle phenotypes, the desired response is an increase in the cell cycle arrest (Red: increased cell cycle arrest. Green: Decreased cell cycle arrest). **F.** The top ten most effective single-agent in silico perturbations for each phenotype. The listed proteomic entities for each phenotype are potential drug targets for slowing down the growth of melanoma cell. The error bars represent the ±SEM over the simulations of all model solutions. For complete prediction heatmaps for phenotypes see Supplementary figure 6-7.

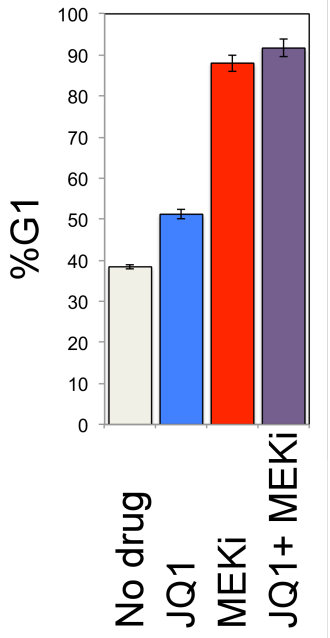
A


cell viability (% no drug ctrl)
CDK4i (RY) [uM]
cell viability (% no drug ctrl)
PKCi (RO) [uM]

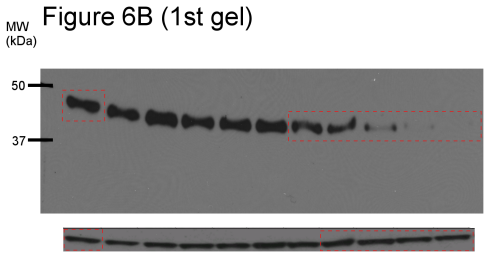
B

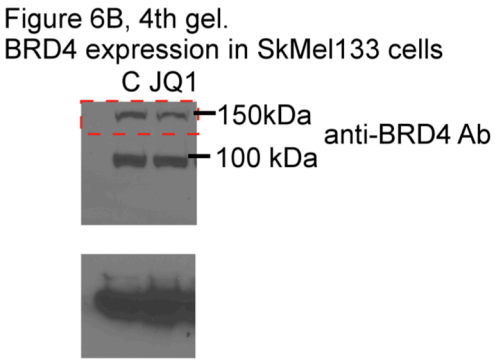

c-Myc
Perturbation Conditions

C


%G1
No drug
JQ1
MEKi
JQ1+ MEKi

D- Uncropped gels





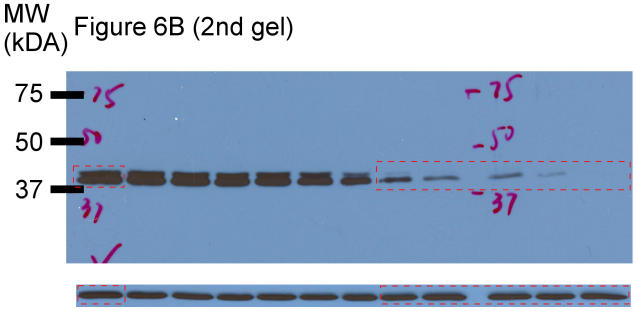
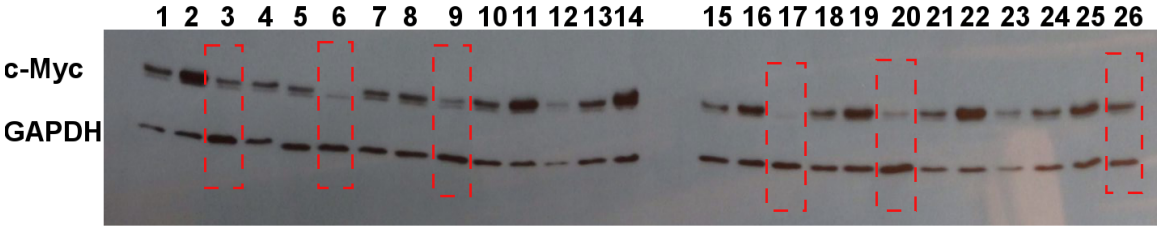


Figure 6B (3rd gel, c-Myc response to JQ1, CDK4i, MEKi and RAFi)



Ab: c-Myc Exposure: 3hrs

Lane No	Drug	t (hr)	Lane no	Drug	t(hr)
1	JQ1	0.5	14	No drug	6
2	JQ1	6	15	JQ1+MEKi	0.5
3	JQ1	20	16	JQ1+MEKi	6
4	MEKi	0.5	17	JQ1+MEKi	20
5	MEKi	6	18	JQ1+RAFi	0.5
6	MEKi	20	19	JQ1+RAFi	6
7	RAFi	0.5	20	JQ1+RAFIi	20
8	RAFi	6	21	JQ1+CDK4i	0.5
9	RAFi	20	22	JQ1+CDK4i	6
10	CDK4i	0.5	23	JQ1+CDK4i	20
11	CDK4i	6	24	No drug	0.5
12	CDK4i	20	25	No drug	6
13	No drug	0.5	26	No drug	20

Figure S5. A. The cell viability response to CDK4i and PKCi. The IC50 for CDK4i is ~2μM in SkMel133 cells. The maximum response is reached ~10μM B. The IC50 for CDK4i is ~1μM and maximum response is reached at 4μM in SkMel133. CDK4i (2μ) and PKCi (3.5-7μM) were used in original input perturbation data. The models recapitulated the effect of high dose PKCi and CDK4i to inhibit cell viability. However, no significant effect on viability is observed in the nM range. **B**. **Changes in c-Myc level in response to perturbations in Skmel133 cells as measured in RPPA experiments.** Each data point is log normalized with respect to the c-Myc level in unperturbed condition. c-Myc level is highest in the presence of CDK4i. Various drug combinations that include STAT3 and mTOR inhibitors follow CDK4i combinations. cMyc level is lowest when cells are perturbed with MEK, BRAF, SRC, PKC and HDAC inhibitors, respectively. Each data-point is the average of RPPA readouts from three replicates. C. **Cell cycle arrest response of melanoma cells to JQ1 and MEKi combination**. 51% and 88% of cells are in G1 stage 24 hours after JQ1 and MEKi treatment, respectively. JQ1+MEKi combination has an increased effect on G1 cell cycle arrest (G1=92%). 39% of cells are in G1 when they are not treated with drugs. Error bars in right panel: ±SE in 3 biological replicates. **D.** Uncropped versions of the gels presented in Figure 6B. In the fourth gel, the observed band ~100kDa is potentially the BRD4 B or C isoform.

Supplementary Tables:

Table S1. Drugs used in perturbation experiments

Drug name	Target	Downstream effector	Dose 1 (nM)	Dose 2 (nM)
AKTi 1/2	AKT	AKTpS473	5.000	10.000
Temsirolimus	mTOR	S6pS240	0.3	0.6
ZSTK474	PI3K	AKTpS473	600	1200
PD0325901	MEK1/2	MAPK1pT202	1.5	3
Stattic	STAT3	STAT3pY705	600	1200
HNHA	HDAC	P27/Kip1	6,000	12,000
PLX4720	$BRAF^{V600E}$	MAPK1pT202	60	120
RO-31-7549	PKC	S6pS240	3500	7000
RY	CDK4	Rb_pS807	2000	NA
P6	JAK	GSK3pS21	20	40
SR	SRC	GSK3pS21	2,400	4,800
Nutlin	MDM2	4EBP1pS65	3,000	6,000

Table S2 Proteins that respond to at least one perturbation and exist in models

Protein	Phosphorylation	Protein	Phosphorylation
4EBP1	S65	IGF1Rβ	
4EBP1	T37	IGFBP2	
4EBP1	T70	IRS1	S307
α -Tubulin		IRS1	
ACC1	S79	ERK/MAPK1	T202
ACC1		MEK	S217
AKT	S473	mTOR	S2448
AKT	T308	p21/Cip1	
AKT		p27/Kip1	
AMPK	T172	p38/MAPK14	T180
ATM	S1981	p38/MAPK14	
ATR		P53	
β-Catenin	S33	p70S6K	T389
β-Catenin		PAI-1	
BAK		PAX2	
BCL-XL		PCNA	
BCL2		PDK1	S241
BIM		PI3Kp85	
BRAF		PKCa	
c-JUN	S73	PLK1	
c-Myc		RAD51	
Caspase9		Rb	S807
Caspase9clvdAsp31		S6	S235
Caveolin		S6	S240
CHK1	S345	S6	
CHK2	T68	SMAD3	S423
Collagenase		SMAD3	
COX2		SRC	Y527
cRAF		SRC	Y416
CyclinB1		SRC	
CyclinD1		STAT3	Y705
CyclinE1		STAT3	
EGFR		STAT5	Y694
ELK1	S383	STAT5	
ER-α		STAT6	Y641
Fibronectin		TAZ	S89
GATA3		TSC2	T1462
GSK3-αβ	S21	TSC2	
GSK3-αβ	S9	XRCC1	
GSK3-αβ		YAP	S127
HSP27		YBI	S102

Table S3. Perturbation conditions

Index	Perturbation1	Dose1 (nM)	Perturbation2	Dose2 (nM)	Index	Perturbation1	Dose1 (nM)	Perturbation2	Dose2 (nM)
1	MEKi	1.5			**46**	BRAFi	60	MDM2i	3000
2	MEKi	1.5	HDACi	6,000	**47**	BRAFi	60	JAKi	20
3	MEKi	1.5	MDM2i	3,000	**48**	BRAFi	60	PKCi	3500
4	MEKi	1.5	JAKi	20	**49**	BRAFi	60	SRCi	2400
5	MEKi	1.5	BRAFi	60	**50**	BRAFi	60	STAT3i	600
6	MEKi	1.5	PKCi	3500	**51**	BRAFi	60	mTORi	0.3
7	MEKi	1.5	SRCi	2400	**52**	PKCi	3500		
8	MEKi	1.5	STAT3i	600	**53**	PKCi	3500	mTORi	0.3
9	MEKi	1.5	mTORi	0.3	**54**	PKCi	7000		
10	MEKi	1.5	PI3Ki	600	**55**	CDK4i	2,000		
11	MEKi	3			**56**	CDK4i	2,000	MEKi	1.5
12	AKTi	10,000			**57**	CDK4i	2,000	AKTi	5000
13	AKTi	5,000			**58**	CDK4i	2,000	HDACi	6000
14	AKTi	5,000	MEKi	1.5	**59**	CDK4i	2,000	MDM2i	3000
15	AKTi	5,000	HDACi	6000	**60**	CDK4i	2,000	JAKi	20
16	AKTi	5,000	MDM2i	3000	**61**	CDK4i	2,000	BRAFi	60
17	AKTi	5,000	JAKi	20	**62**	CDK4i	2,000	PKCi	3500
18	AKTi	5,000	BRAFi	60	**63**	CDK4i	2,000	SRCi	2400
19	AKTi	5,000	PKCi	3500	**64**	CDK4i	2,000	STAT3i	600
20	AKTi	5,000	SRCi	2400	**65**	CDK4i	2,000	mTORi	0.3
21	AKTi	5,000	STAT3i	600	**66**	CDK4i	2,000	PI3Ki	600
22	AKTi	5,000	mTORi	0.3	**67**	SRCi	2,400		
23	AKTi	5,000	PI3Ki	600	**68**	SRCi	2,400	PKCi	3500
24	HDACi	12,000			**69**	SRCi	2,400	mTORi	0.3
25	HDACi	6,000			**70**	SRCi	4,800		
26	HDACi	6,000	MDM2i	3000	**71**	STAT3i	600		
27	HDACi	6,000	PKCi	3500	**72**	STAT3i	600	HDACi	6000
28	HDACi	6,000	SRCi	2400	**73**	STAT3i	600	MDM2i	3000
29	HDACi	6,000	mTORi	0.3	**74**	STAT3i	600	PKCi	3500
30	MDM2i	3,000			**75**	STAT3i	600	SRCi	2400
31	MDM2i	3000	PKCi	3500	**76**	STAT3i	600	mTORi	0.3
32	MDM2i	3000	SRCi	2400	**77**	STAT3i	1200		
33	MDM2i	3000	mTORi	0.3	**78**	mTORi	0.3		
34	MDM2i	6,000			**79**	mTORi	0.6		
35	JAKi	20			**80**	PI3Ki	600		
36	JAKi	20	HDACi	6000	**81**	PI3Ki	600	HDACi	6000
37	JAKi	20	MDM2i	3000	**82**	PI3Ki	600	MDM2i	3000
38	JAKi	20	PKCi	3500	**83**	PI3Ki	600	JAKi	20
39	JAKi	20	SRCi	2400	**84**	PI3Ki	600	BRAFi	60
40	JAKi	20	STAT3i	600	**85**	PI3Ki	600	PKCi	3500
41	JAKi	20	mTORi	0.3	**86**	PI3Ki	600	SRCi	2400
42	JAKi	40			**87**	PI3Ki	600	STAT3i	600
43	BRAFi	120			**88**	PI3Ki	600	mTORi	0.3
44	BRAFi	60			**89**	PI3Ki	1,200		

Table S4. Predicted G1-arrest response to combined targeting of c-Myc and other nodes (Perturbations on c-Myc + X). The response values are in log2 and normalized with respect to the response to single agent c-Myc perturbation ($G1_{c\text{-}Myc+X}/G1_{c\text{-}Myc}$). The top 20 paired perturbation conditions are ranked based on the response to the lowest dose (c-Myc_IC20). The _ICX tag determines the strength of the in silico perturbation. The X (e.g, 20, 40) represents the % decrease in the target node activity at the given perturbation strength.

	c-Myc_IC20	c-Myc_IC40	c-Myc_IC60	c-Myc_IC80
aMEK_IC80	2.316	1.971	1.870	1.835
aMEK_IC60	2.253	1.924	1.824	1.791
aMEK_IC40	2.071	1.786	1.698	1.668
CyclinD1_IC80	1.928	1.699	1.630	1.607
CyclinD1_IC60	1.893	1.670	1.606	1.581
aBRAFm_IC80	1.889	1.677	1.611	1.586
aBRAFm_IC60	1.845	1.640	1.575	1.554
CyclinD1_IC40	1.795	1.594	1.536	1.515
MAPKpT202_IC80	1.725	1.535	1.478	1.458
aBRAFm_IC40	1.721	1.540	1.482	1.463
aSRC_IC80	1.694	1.541	1.488	1.471
MAPKpT202_IC60	1.686	1.506	1.453	1.434
aMEK_IC20	1.682	1.493	1.434	1.419
PLK1_IC80	1.679	1.531	1.479	1.462
aSRC_IC60	1.678	1.526	1.474	1.456
PLK1_IC60	1.670	1.523	1.470	1.452
MEKpS217_IC80	1.658	1.482	1.431	1.416
PLK1_IC40	1.639	1.493	1.445	1.430
aSRC_IC40	1.621	1.473	1.425	1.408